



King's Research Portal

Document Version
Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Catmur, C., Press, C., & Heyes, C. (2016). Mirror Neurons from Associative Learning. In R. Murphy, & R. Honey (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Learning* WILEY-BLACKWELL.

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Mirror Neurons from Associative Learning

Caroline Catmur¹, Clare Press² & Cecilia Heyes^{3,4*}

¹Department of Psychology, University of Surrey

²Department of Psychological Sciences, Birkbeck, University of London

³All Souls College, University of Oxford

⁴Department of Experimental Psychology, University of Oxford

*Corresponding author: cecilia.heyes@all-souls.ox.ac.uk

Abstract:

Mirror neurons fire both when executing actions and observing others perform similar actions. Their sensorimotor matching properties have generally been considered a genetic adaptation for social cognition; however, in the present chapter we argue that the evidence in favor of this account is not compelling. Instead we present evidence supporting an alternative account: that mirror neurons' matching properties arise from associative learning during individual development. Notably, this process was not 'designed' by genetic evolution specifically to produce mirror neurons, but just happens to produce them when the developing system receives correlated experience of observing and executing similar actions. Sensorimotor experience with non-matching actions, or with objects and actions, is hypothesized to generate other cell types in the same regions through the same process. The associative account has major implications for research into mirror neuron function and suggests several important lines of future research.

Keywords: Mirror neurons, associative learning, sensorimotor, development, imitation, action understanding, adaptation, mirror system, fMRI, TMS

Associative learning theory has typically been used to explain the behavior of whole animals; to understand why organisms make particular kinds of responses to focal stimuli and contextual cues. In this chapter, we use research on associative learning in a slightly different way, in an attempt to explain the behavior – the firing patterns – of individual neurons, rather than whole animals. The neurons in question are known as ‘mirror neurons’ (MNs), and the behavior that has made mirror neurons famous is their tendency to fire not only when a macaque performs an action, but also when the macaque passively observes a similar action performed by another. Neurons with this capacity to match observed and executed actions were originally found in area F5 of the ventral premotor cortex (PMC; di Pellegrino et al., 1992) and subsequently in the inferior parietal lobule (IPL; Fogassi et al., 2005) of the macaque brain. A substantial body of evidence now suggests that MNs are also present in the human brain (Molenberghs, Cunnington, & Mattingley, 2012).

A variety of functions have been ascribed to MNs. Popular suggestions relate to action understanding (Gallese & Sinigaglia, 2011; Rizzolatti et al., 1996), imitation (Iacoboni et al., 1999) and language processing (Rizzolatti & Arbib, 1998). A great deal of interest has also been generated in the wider scientific and public media: MNs have been hailed as “cells that read minds” (Blakesee, 2006), “the neurons that shaped civilization” (Ramachandran, 2009), and a “revolution” in understanding social behavior (Iacoboni, 2008).

Whereas much research has focused on theorizing and speculation about MN functions, this chapter’s primary focus is the origin of MNs. We ask not ‘What are MNs for?’, but ‘What is the process that gives MNs their ‘mirroriness’; their fascinating capacity to match observed with executed actions?’ The standard answer to this question (e.g. Rizzolatti & Craighero, 2004) is evolution. The ‘adaptation account’ assumes that the mirroriness of MNs was produced by natural selection acting on genetic variation. In contrast, we will argue that the balance of evidence supports the ‘associative account’ (Catmur et al., in press; Cook et al., in press; Heyes, 2010); it

suggests that the mirroriness of MNs is produced in the course of individual development by sensorimotor associative learning. We will also argue that the associative model has major methodological implications for research investigating the functions of MNs.

The first section outlines key background information regarding MNs in macaques and humans. Next we present the adaptation and associative accounts. In the following section we introduce four kinds of evidence that have the potential to favor one of these hypotheses over the other, and discuss each of these types of evidence in turn. Finally we examine the implications of the associative account for future research investigating the functions of MNs.

<A>Mirror neuron background information

Where are they found, and what qualifies as a mirror neuron?

In the macaque, ‘classical’ MN areas include ventral PMC and IPL (see Figure 1). However, MNs have also been found in ‘non-classical’ areas, including primary motor cortex and dorsal PMC (Dushanova & Donoghue, 2010; Tkach, Reimer, & Hatsopoulos, 2007). In humans, there is evidence at both the single-cell and population level of neurons with sensorimotor matching properties. These have been found both in ‘classical’ MN areas, including inferior frontal gyrus (IFG; considered the human homologue of macaque F5) (Kilner et al., 2009) and inferior parietal cortex (Chong et al., 2008), and in non-classical areas, including dorsal PMC, superior parietal lobule, and cerebellum (Molenberghs et al., 2012), occipitotemporal cortex (Oosterhof, Tipper, & Downing, 2012), supplementary motor area, and medial temporal lobe (Mukamel et al., 2010).

[Figure 1 near here]

Although some researchers only refer to neurons found in classical areas as MNs (e.g. Molenberghs et al., 2012), many others, like us, use the term ‘MN’ to refer to neurons in both classical and non-classical areas (Gallese & Sinigaglia, 2011; Keysers & Gazzola, 2010). Functional definitions of what constitutes a MN also vary. In some cases, the term ‘MN’ is used to refer to any neuron that fires during both the execution and observation of action, regardless of whether the executed and observed actions are similar to one another (Gallese et al., 1996; Rizzolatti & Craighero, 2004). In contrast, and following the majority of researchers in the field, we consider that MNs’ ‘mirrorness’ is defined by the fact that they respond to observation and execution of similar actions. However, following common usage, we also refer to ‘logically-related’ MNs (see following subsection), which fire during observation and execution of dissimilar actions which have some functional relation, as ‘MNs’.

Mirror neuron response properties in the macaque

Macaque MNs have been broadly divided into three types (Figure 2), based on their field properties—the sensory and motoric conditions in which they fire: ‘Strictly congruent’ MNs discharge during observation and execution of the same action, for example, a ‘precision’ grip made with thumb and index finger. ‘Broadly congruent’ MNs fire during the execution of one action (e.g., precision grip) and during the observation of one or more similar, but not identical, actions (e.g., only power grip; or multiple actions e.g. precision grip, power grip, *and* grasping with the mouth). So-called ‘logically-related’ MNs (di Pellegrino et al., 1992) respond to different actions in observe and execute conditions. For example, they fire during the observation of an experimenter placing food in front of the monkey, and when the monkey executes a grasp on the food in order to eat it (it is likely that cells with these properties were dubbed ‘logically-related’, not because there is a formal relationship between their eliciting conditions, but to acknowledge that, unlike other MNs, they do not match or ‘mirror’ observed and executed actions). MNs do not respond to the presentation of objects alone

(di Pellegrino et al., 1992). However, ‘canonical neurons’, which are active during object observation and also during execution of an action that is commonly performed on that object, are located alongside MNs in both premotor and parietal areas (Murata et al., 1997, 2000).

[Figure 2 near here]

Macaque MNs fire during execution and observation of a broad range of hand and mouth actions. The hand actions include grasping, placing, manipulating with the fingers and holding (di Pellegrino et al., 1992). The mouth actions include ingestive behaviors such as breaking food items, chewing and sucking, and communicative gestures such as lip-smacking, lip-protrusion and tongue-protrusion (Ferrari et al., 2003).

Mirror neurons in humans

Only one study offers single cell recording evidence of MNs in the human brain (Mukamel et al., 2010). However, a considerable body of evidence from neuroimaging, transcranial magnetic stimulation (TMS), and behavioral studies, summarized in the following subsections, suggests that human brains contain MNs or comparable ‘mirror mechanisms’ (Glenberg, 2011; referred to throughout this chapter as MNs).

Functional magnetic resonance imaging (fMRI) has identified regions of PMC and inferior parietal areas that respond during both action observation and execution (Gazzola & Keysers, 2009; Iacoboni et al., 1999; Vogt et al., 2007). More recently, ‘repetition suppression’ effects, whereby the neural response reduces when events activating the same neuronal population are repeated, (Grill-Spector, Henson, & Martin, 2006) provide further evidence for the presence of ‘mirror’ neuronal populations. Action observation followed by execution of the same action, or vice versa, elicits a suppressed

response in inferior parietal regions (Chong et al., 2008; Lingnau, Gesierich, & Caramazza, 2009) and in PMC (Kilner et al., 2009; Lingnau et al., 2009), indicating that the same neuronal population is active when observing and executing the same action. Multivariate pattern analysis (MVPA) has also revealed cross-modal action-specific representations consistent with the presence of ‘mirror’ neuronal populations (Oosterhof et al., 2012): a ‘classifier’ program trained to discriminate neural responses to the execution of different actions can subsequently, when tested with neural responses to the observation of those actions, detect which action was observed, suggesting that the same neural representations encode action observation and execution.

‘Mirror’ patterns of motor-evoked potentials (MEPs) further suggest a human mirror mechanism (Fadiga et al., 1995). When TMS is applied to M1 during passive action observation, the amplitude of MEPs recorded from the muscles required to execute that action increases. For example, observing index and little finger movements selectively facilitates the amplitude of MEPs recorded from the muscles responsible for index and little finger movements (Catmur et al., 2011). That action observation selectively increases corticospinal excitability to action relevant muscles is suggestive of ‘mirror’ sensorimotor connectivity.

Behaviorally, automatic imitation occurs when observation of an action involuntarily facilitates performance of a topographically similar action and/or interferes with performance of a topographically dissimilar action (Brass, Bekkering, & Prinz, 2001; Stürmer, Aschersleben, & Prinz, 2000). Humans show robust automatic imitation when they observe hand, arm, foot and mouth movements (Heyes, 2011). This is regarded by many researchers as evidence of a human mirror mechanism (Ferrari, Bonini, & Fogassi, 2009; Iacoboni, 2009; Kilner, Paulignan, & Blakemore, 2003).

<A>Mirror neurons’ ‘mirrorness’: Adaptation or association?

Here we outline the standard, adaptation account of the origin of MNs, and the alternative associative account. Both accounts assume that genetic information and experience contribute to the development of MNs. They differ in the roles they assign to genetic evolution and to learning in producing MNs' characteristic matching properties.

The adaptation account suggests that the matching properties of MNs are an adaptation for action understanding and/or related social cognitive abilities (the term 'adaptation' is used here to describe a phenotypic characteristic that is genetically inherited, and that was favored by natural selection to fulfill a particular function or 'purpose'; Williams, 1966). Specifically, the adaptation account assumes that among common ancestors of macaques and humans, some individuals had a stronger genetic predisposition to develop MNs with matching properties, and that these individuals were more reproductively successful than those with a weaker genetic predisposition because the development of MNs enhanced their capacity to understand others' actions. Consequently, a genetic predisposition to develop MNs became universal, or nearly universal, in macaques and humans. The adaptation account further suggests that motor experience (executing actions) and/or visual experience (observing actions) plays a facilitative or 'triggering' (Gottlieb, 1976; Ariew, 2006) role in the development of MNs, but their 'mirror', sensory-to-motor matching properties are due to this genetic predisposition.

The adaptation account has largely been set out in discussions of the 'evolution' of MNs (Gallese & Goldman, 1998; Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004; Rochat et al., 2010). For example, it was suggested that "the mirror neuron mechanism is a mechanism of great evolutionary importance through which primates understand actions done by their conspecifics" (Rizzolatti & Craighero, 2004, p. 172). A number of discussions have also suggested that MNs are present at birth (Ferrari et al., 2009; Gallese et al., 2009; Lepage & Theoret, 2007; Rizzolatti & Fadiga, 1998), a feature commonly associated with adaptations (Mameli & Bateson, 2006).

In contrast, the associative account suggests that the matching properties of MNs are not a product of a specific genetic predisposition, but instead result from domain-general processes of associative learning (Catmur et al., in press; Cook et al., in press; Heyes, 2010). Associative learning is found in a wide range of vertebrate and invertebrate species, indicating that it is an evolutionarily ancient and highly conserved adaptation for tracking predictive relationships between events (Heyes, 2012; Schultz & Dickinson, 2000).

[Figure 3 near here]

Figure 3 represents a theory (Heyes & Ray, 2000; Heyes, 2010) of how MNs might acquire their matching properties through sensorimotor associative learning. Before associative learning, sensory neurons responsive to different high-level visual properties of observed action (Oram & Perrett, 1994, 1996) are weakly connected, directly or indirectly, to motor neurons in parietal cortex (Gallese et al., 2002) and PMC (Rizzolatti et al., 1988). Although some of these connections may be stronger than others, the links between sensory and motor neurons coding similar actions are not consistently stronger than other, nonmatching links. Correlated (i.e. contiguous and contingent) excitation of sensory and motor neurons that code similar actions produces MNs. For example, when an adult imitates an infant's facial movements, there might be correlated excitation of neurons that are responsive to the observation and execution of lip protrusion. Correlated excitation of the sensory and motor neurons increases the strength of the connection between them, so that subsequent excitation of the sensory neuron propagates to the motor neuron. Thereafter, the motor neuron fires, not only during execution of lip protrusion, but also during observation of lip protrusion, via its connection with the sensory neuron; what was originally a motor neuron has become a lip protrusion MN. In humans, there are many possible sources of correlated excitation of sensory and motor neurons encoding the same action. It occurs not only

when we are imitated, but also when we observe our own actions – directly or using an optical mirror; observe others during synchronous activities – e.g. in sports and dance training; and via ‘acquired equivalence’ experience, for example, when the same sound (a word, or a sound produced by an action, e.g. lip-smacking) is paired sometimes with observation of an action and sometimes with its execution (Ray & Heyes, 2011). In all of these situations, motor activity is not initiated by, but it is correlated with, observation of matching actions.

Thus the associative account identifies sources in everyday life of the kind of correlated sensorimotor experience necessary for MN development, and many of these sources are sociocultural; to a large extent, MNs are built through social interaction. Another important point to note about the associative account is its emphasis on contingency. Following contemporary associative learning theory, it anticipates that the mature properties of MNs will covary, not only with the number of occasions on which observation of an action has been paired with its execution (contiguity), but also, as a result of context blocking, with the relative predictiveness of observation for execution, or vice versa (contingency; Cook et al., 2010). Experiments testing the associative account are discussed below.

In summary: the associative account implies that the characteristic, matching properties of MNs result from a genetically evolved process, associative learning, but that this process was not ‘designed’ by genetic evolution specifically to produce matching MNs. It just happens to produce matching MNs when the developing system receives correlated experience of observing and executing similar actions. When the system receives correlated experience of observing objects and executing actions, the same associative process produces canonical neurons. When the system receives correlated experience of observing one action and executing a different action, the same associative process produces logically-related MNs.

Thus, the adaptation account says that genetic evolution has played a specific and decisive role, and learning plays a merely facilitative role, in the development of matching MNs. In contrast, the associative account says that evolution has played a non-specific background role, and that the characteristic matching properties of MNs are forged or ‘induced’ (Gottlieb, 1976) by sensorimotor learning.

<A>Distinguishing the adaptation and associative accounts

Here we present the four evidence-based arguments which aid in distinguishing between the adaptation and associative accounts. The first argument provides the foundation for the adaptation account. It suggests that examination of the field properties of MNs – and, in particular, their ‘goal’ coding – forces the conclusion that MNs are ‘designed’ (Williams, 1966) for action understanding. In the following subsection we examine the field properties of MNs and suggest that this argument is not compelling.

The second argument suggests that research using conditioning procedures shows associative learning to be the right kind of learning to produce MNs. Specifically, the ways in which associative learning tracks contingent relationships, and enables contextual modulation of these connections, makes it apt to produce MNs (and non-matching visuomotor neurons) in typical developmental environments.

We then draw on research examining the development of MNs and their modification through sensorimotor experience. First we discuss research with infants and adults that has been used to support a “poverty of the stimulus” argument (Chomsky, 1975); to suggest that MNs emerge too early in development, after too little sensorimotor experience, to have been forged by associative learning. In contrast, we offer a “wealth of the stimulus” argument.

Finally, we focus on evidence that, even in adulthood, the properties of MNs can be changed in radical ways by relatively brief periods of sensorimotor experience. This evidence supports the associative account in two ways: it confirms novel predictions of the associative account, and indicates that the development of MNs is not buffered or protected from perturbation in the way one would expect if MNs were an adaptation for action understanding.

Do mirror neurons encode the ‘goal’ of an action?

Supporters of the adaptation account (e.g., Rizzolatti & Sinigaglia, 2010) argue that examination of the field properties of MNs indicates that they encode ‘goals’. They further argue that this property suggests that MNs evolved to mediate action understanding. We first, therefore, consider how well the neurophysiological data accord with this view. The term ‘goal’ has numerous interpretations (Hickok, 2009). We will consider two commonly adopted definitions, assuming that MNs encode ‘goals’ if they encode (1) object-directed actions or (2) high-level action intentions.

Early descriptions of MN field properties reported that intransitive, that is non-object-directed, actions (e.g. tongue-protrusion) and pantomimed actions (e.g. miming a precision grip without an object) did not elicit MN responses (di Pellegrino et al., 1992; Gallese et al., 1996). In contrast, robust responses were reported when monkeys observed object-directed actions. This pattern raises the possibility that MNs encode ‘goals’ in the sense that they are responsive only to object-directed actions. However, a close reading of the single cell data suggests that only a small subset of MNs appear to encode action goals in these terms. A subset of the MNs described in the early reports continued to respond, albeit less strongly, to pantomimed or intransitive actions (di Pellegrino et al., 1992; Gallese et al., 1996 Figure 5b). Subsequent studies confirmed that sizable proportions, perhaps the majority, of MNs exhibit robust responses to the observation of object-free body

movements, such as lip-smacking, lip-protrusion and tongue-protrusion (Ferrari et al., 2003). Also, as reported by Kraskov et al. (2009), 73% of MN responses modulated by observation of object-directed grasping showed similar modulation during observation of pantomimed grasping.

As well as referring to the object of an action, the term 'goal' has also been used to refer to what the actor intends to achieve – for example, “grasp in order to eat” (Fogassi et al., 2005) or “taking possession of an object” (Rochat et al., 2010). Rizzolatti and Sinigaglia (2010, p. 269) state: “only those [neurons] that can encode the goal of the motor behaviour of another individual with the greatest degree of generality can be considered to be crucial for action understanding”. The suggestion that MNs encode high-level action intentions is consistent with reports that some broadly congruent MNs respond to the observation of multiple actions; for example any 'grasping' action executed with the hand or mouth (Gallese et al., 1996). It is also made plausible by reports that MN responses to grasping can be modulated by the final outcome of the motor sequence (Bonini et al., 2010; Fogassi et al., 2005).

However, the single cell data again suggest that relatively few MNs have the field properties one would expect of a system designed to represent high-level action intentions. For example, Gallese et al. (1996) reported that during action observation 37.5% of MNs responded differently depending on whether the action was executed with the left or right hand, and 64% showed direction sensitivity, preferring either left-to-right or right-to-left grasping actions. Similarly, many MNs (53%) respond selectively to the observation of actions executed within ('peripersonal' MNs) or beyond ('extrapersonal' MNs), not the actor's, but the observing monkey's reach (Caggiano et al., 2009). The majority (74%) of MNs also exhibit view-dependent responses; some MNs are tuned to egocentric (first-person) presentation, while others respond maximally to allocentric (third-person) perspectives (Caggiano et al., 2011). Each of these classes of MN is sensitive to features of action that fall well below the “greatest degree of generality”, and of intentions such as “grasping in order

to eat” or “taking possession of an object”.

Associative learning: the right kind of learning to generate mirror neuron field properties?

The previous subsection suggested that many MNs have field properties incompatible with the hypothesis that they were designed by evolution to mediate action understanding via goal coding. Here, in complementary fashion, we argue that research on the roles of contingency and contextual modulation in associative learning enables the associative account to provide a unified explanation of all MN field properties reported to date.

Associative learning depends not only on contiguity – events occurring close together in space and time – but also on contingency: the degree to which one event reliably predicts the other (Elsner & Hommel, 2004; Rescorla, 1968; Schultz & Dickinson, 2000). The associative account therefore anticipates that MNs will acquire sensorimotor matching properties only when an individual experiences systematic contingencies between sensory events and performed actions (Cooper et al., 2013). This feature of associative learning ensures that the matching properties of MNs reflect sensorimotor relationships that occur reliably in the individual’s environment, rather than chance co-occurrences. Cook and colleagues (2010) described evidence that the human mirror mechanism is modified by contingent but not by non-contingent sensorimotor experience.

Sensitivity to contingency explains the mix of strictly-congruent MNs, sensitive to the low-level features of observed actions (type of grip, effector used, direction of movement, view-point, proximity to the observer), and broadly-congruent MNs, responsive to multiple related actions irrespective of the manner of their execution. Both visual and motor systems are known to be organized hierarchically (Jeannerod, 1994; Perrett et al., 1989), comprising different populations encoding relatively low-level (e.g., descriptions of particular ‘precision’ or ‘power’ grips) and more

abstract representations (e.g., descriptions of ‘grasping’). Crucially, contingencies can be experienced between both low- and high-level sensory and motor representations. When a monkey observes itself performing a precision grip, the excitation of sensory and motor populations encoding a specific grip (low-level) are correlated. However, during group feeding, a monkey might observe and perform a range of grasping actions, thereby causing correlated excitation of higher-level visual and motoric descriptions of grasping. Contingency sensitivity therefore explains the existence of both strictly-congruent MNs, tuned to a particular sensory representation (e.g. a right-to-left precision grip executed with the right hand viewed allocentrically in extrapersonal space), and broadly congruent MNs, responsive to the observation of a number of related actions (see Figure 4).

[Figure 4 near here]

Contingency sensitivity also explains other MN properties. According to the associative account, MNs acquire sensorimotor properties whenever individuals experience a contingency between ‘seeing’ and ‘doing’. Crucially, there is no requirement that contingencies be between action execution and observation of the same action. Both monkeys and humans frequently experience non-matching sensorimotor contingencies, where the observation of one action predicts the execution of another; for example, you release and I grasp (Newman-Norlund et al., 2007). The associative account therefore explains the existence of logically-related MNs that respond to different actions in observe and execute conditions. Equally, there is no requirement that contingencies be between action execution and the perception of ‘natural’ action-related stimuli, such as the sight of animate motion or sounds that could have been heard by ancestors of contemporary monkeys. Thus the associative account explains why ‘tool-use’ MNs (Ferrari, Rozzi, & Fogassi, 2005) develop when action execution (e.g. grasping a food item) is reliably predicted by the sight of actions performed with tools (e.g. seeing food items being gripped with pliers), and why ‘audiovisual’ MNs (Keysers et al., 2003; Kohler et al., 2002) develop when action performance

predicts characteristic action sounds (e.g. paper tearing or plastic crumpling) (Cook, 2012): there is a high contingency between the sight of the experimenter gripping food with pliers and the subsequent execution of a grasp by the macaque; and between the sound of paper tearing and the execution of the ripping action that produces that sound.

Studies of conditioning that have supported the role of contingency indicate that learned responses acquired under contingency control are often also subject to contextual control; if a stimulus is associated with two responses, each in a different context, then the context determines which association, representing a response-outcome contingency, is cued by the stimulus (Bouton, 1993, 1994; Peck & Bouton, 1990). For example, Peck and Bouton (1990) initially placed rats in a conditioning chamber with a distinctive scent (e.g., coconut) where they learned to expect electric shock following a tone. The rats were then transferred to a second chamber with a different scent (e.g., aniseed) where the same tone predicted the delivery of food. The rats quickly learned the new contingency and conditioned foraging responses replaced conditioned freezing. However, learning in the second phase was context dependent. When returned to the first chamber, or transferred to a third chamber with a novel scent, the tone once again elicited freezing. The associative account of MN properties draws on the components of associative learning theory that explain this kind of effect.

Using associative learning theory in this way, several findings from the MN literature can be interpreted in terms of contextual modulation of MN firing (Cook, Dickinson, & Heyes, 2012). For example, some MNs show stronger visual responses to object-directed grasping than to pantomimed grasping in object-absent contexts (Gallese et al., 1996), and in some cases, the modulating influence of the object-context can be seen even when the target object is occluded prior to contact with the hand (Umiltà et al., 2001). Similarly, MN responses during the observation of grasping may be modulated by the type of object being grasped (Caggiano et al., 2012), with some

MNs responding strongly in the presence of high-value objects (food, non-food objects predictive of reward), and some in the presence of low-value objects (non-food objects not associated with reward). In the clearest example, the same motor act, grasping with a precision grip, elicits different MN responses dependent on whether the action is observed in the presence (“grasp to place”) or absence (“grasp to eat”) of a plastic cup (Bonini et al., 2010; Fogassi et al., 2005). Rather than the plastic cup providing a cue to the actor’s intention, it may act as an associatively generated contextual cue modulating the operation of two associations. In the same way that the sound of the tone elicited different behaviors when presented in the coconut and aniseed contexts (Peck & Bouton, 1990), observing a precision grip may excite different MNs in the cup-present and cup-absent contexts (see Figure 4). Thus, while many of the field properties described above are frequently cited as evidence of goal (intention) coding by MNs, they are equally consistent with contextual modulation within an associative framework.

Sufficient opportunity for learning before mirror neurons emerge?

MNs have not been measured directly in neonates. However, other research involving infants has been used to support a “poverty of the stimulus” (Chomsky, 1975) argument suggesting that MNs emerge too early in development, after too little sensorimotor experience, to have been forged by associative learning. Specifically, it has been claimed that imitation is mediated by MNs, and that both human and macaque infants are able to imitate when they have had minimal opportunity for sensorimotor learning. However, the evidence supporting the second claim is not compelling. Building on previous analyses (e.g. Anisfeld, 1996), a recent review found evidence that human neonates ‘copy’ only one action – tongue protrusion – and that, since tongue protrusion occurs in response to a range of arousing stimuli, this ‘copying’ does not show the specificity that is characteristic of imitation or of MNs (Ray & Heyes, 2011).

Turning to macaque infants, Ferrari et al. (2006) reported immediate imitation of tongue protrusion and lip-smacking in 3 day-old macaques. However, the effects were not present on days 1, 7 and 14 postpartum, and it is not clear whether they were replicated in a subsequent study (Paukner, Ferrari, & Suomi, 2011). The later study did report imitation of lip-smacking in macaques less than one week old, but this effect seems to have been due to a low frequency of lip-smacking in the control condition, rather than to an elevated frequency of lip-smacking when the infants were observing lip-smacking. Therefore, in common with the data from human infants, studies of imitation in newborn macaques do not currently support the conclusion that infants can imitate before they have had the opportunity for relevant sensorimotor learning.

A related argument has suggested that the associative account must be wrong because suppression of electroencephalographic (EEG) activity in the alpha frequency range (~6-13 Hz) during action observation (and execution) reflects the operation of MNs; and that both human and macaque infants show alpha suppression when they have had minimal opportunity for sensorimotor learning. In this case, both of the claims are weak. Alpha suppression is found over central cortical regions when observing and executing actions, but it may not reflect the activity of MNs. First, the functional significance of lower band EEG activity is poorly understood even in adults, and is yet more difficult to interpret in infants where, for example, less information is available about the source (Marshall & Meltzoff, 2011). Alpha suppression in other locations is interpreted differently (e.g. as evidence of increased visual processing) and the only neonatal action observation study (Ferrari et al., 2012) has insufficient spatial resolution to provide source information. Second, adult studies have traced the likely source of alpha suppression during action execution to the somatosensory cortex (Hari & Salmelin, 1997), suggesting that alpha suppression during action observation may not index motor processing (and thus MNs) at all (Coll et al., submitted). Third, even if alpha suppression does index motor processing, it does not show that the motor activation matches or mirrors the observed actions (Marshall & Meltzoff, 2011). Thus, alpha suppression during observation of lip-smacking,

which has been reported in neonatal monkeys (Ferrari et al., 2012), may reflect a generalized readiness to act, or motor activation of tongue protrusion or hand movement, rather than motor activation of lip-smacking. Furthermore, it has not been shown that alpha suppression occurs when infants have had insufficient correlated sensorimotor experience to build MNs through associative learning. Indeed, studies of human infants suggest an age-related trend consistent with the associative account: for example, Nyström (2008) found no evidence of alpha suppression when 6-month-old infants observed actions, but effects have been obtained at nine and 14 months (Marshall, Young, & Meltzoff, 2011; Southgate et al., 2010).

It is important to note that although MN activity in newborns would be inconsistent with the associative model, the associative account is predicated on a 'wealth of the stimulus' argument, and therefore anticipates MN activity in young infants following sufficient correlated sensorimotor experience (Ray & Heyes, 2011). This 'wealth argument' points out that typical human developmental environments contain multiple sources of the kind of correlated sensorimotor experience necessary to build MNs; that each of these sources is rich; and that the mechanisms of associative learning can make swift and efficient use of these sources. The range of sources available to young human infants includes self-observation, being imitated by adults, being rewarded by adults for imitation, and acquired equivalence experience in which, for example, the infant hears the same tapping sound when she hits an object herself and when she sees the object hit by another person. A common misconception about associative learning is that it always occurs slowly. On the contrary, when contingency is high, infants can learn action-effect associations in just a few trials (Paulus et al., 2012; Verschuur et al., 2010) and human adults demonstrate rapid learning even with complex contingencies (e.g., Baker, Vallée-Tourangeau, & Murphy, 2000).

The influence of sensorimotor learning

The associative account has been explicitly tested in experiments examining the effects of laboratory-based sensorimotor training on MNs in human adults. Building on the results of more naturalistic studies (Calvo-Merino et al., 2005, 2006; Ferrari et al., 2005; Vogt et al., 2007), these experiments have isolated the effects of sensorimotor experience from those of purely visual and purely motor experience. Using all the measures of MN activity commonly applied to humans (imitation, motor evoked potentials, and fMRI measures including repetition suppression), they have shown that relatively brief periods of sensorimotor experience can enhance (Press, Gillmeister, & Heyes, 2007), abolish (Cook et al., 2012; Cook et al., 2010; Gillmeister et al., 2008; Heyes et al., 2005; Wiggett et al., 2011), reverse (Catmur et al., 2008; Catmur et al., 2011; Catmur, Walsh, & Heyes, 2007; Cavallo et al., 2013) and induce (Landmann et al., 2011; Petroni, Baguear, & Della-Maggiore, 2010; Press et al., 2012) MN activity (details below). These findings reveal the kind of flexibility one would expect if MNs are forged by sensorimotor associative learning. In contrast, this kind of flexibility is hard to reconcile with the adaptation account. If MNs were a genetic adaptation, one would expect their development to be protected or ‘buffered’ against environmental perturbations that were occurring when MNs evolved and that could interfere with their adaptive function (Cosmides & Tooby, 1994; Pinker, 1997). Thus, if MNs are indeed an adaptation for ‘action understanding’, their development should be buffered to prevent them from coding stimulus-response and response-outcome relationships that could interfere with that function. For example, MNs should be prevented from coding inanimate, rather than action, stimuli; and from coding dissimilar, rather than similar, observed and executed actions.

Evidence that MNs are not resistant to coding inanimate stimuli comes from studies showing that arbitrary sound, color and shape stimuli can induce mirror motor evoked potentials (D’Ausilio et al., 2006; Petroni et al., 2010), fMRI responses (Landmann et al., 2011; Press et al., 2012), and behavioral effects (Press et al., 2007) following sensorimotor training. For example, Press and colleagues (2007) gave participants approximately 50 minutes of sensorimotor training in which they

repeatedly opened their hand when seeing a robotic pincer open, and closed their hand when seeing the robotic pincer close. Prior to this training, the pincer movement elicited less automatic imitation than human hand movement, but 24 hours after training, the automatic imitation effect was as strong for the pincer movement as for the human hand.

Evidence that MNs are not resistant to coding dissimilar actions comes from studies showing that non-matching (or 'counter-mirror') sensorimotor training abolishes automatic imitation (Cook et al., 2012; Cook et al., 2010; Gillmeister et al., 2008; Heyes et al., 2005; Wiggett et al., 2011), and reverses both fMRI (Catmur et al., 2008) and MEP mirror responses (Catmur et al., 2007). For example, Catmur and colleagues (2007) gave participants approximately 90 minutes of non-matching sensorimotor training in which they repeatedly made an index finger movement while observing a little finger movement, and vice versa. Before this training they showed mirror MEP responses. For example, observation of index finger movement elicited more activity in an index finger muscle than observation of little finger movement, and vice versa for the little finger muscle. After training, this pattern was reversed. For example, observation of index finger movement elicited more activity in the little finger muscle than observation of little finger movement. Similarly, following sensorimotor training in which observation of hand actions was paired with execution of foot actions and vice versa, fMRI responses to action observation were reversed: premotor and parietal areas normally more responsive to the sight of hand actions now showed stronger responses to observation of foot actions (Catmur et al., 2008).

Thus, a substantial body of evidence from studies of training and expertise has confirmed the predictions of the associative account, showing that mirror responses can be changed in radical ways by sensorimotor learning. In particular, these studies suggest that MNs are not buffered or protected against sensorimotor experience of a kind that makes them code inanimate stimuli and dissimilar actions.

<A>Investigating the contribution of mirror neurons to social behavior

The associative account suggests that MNs do not have a specific biological purpose or 'adaptive function', distinct from that of other neurons with visuomotor properties. However, the associative account leaves open the possibility that MNs are recruited in the course of development to contribute to one or more 'psychological functions'. They could be useful – possibly they could contribute to a variety of social functions – without having been designed by evolution for a particular use. Thus the associative account is functionally permissive; however, it implies that a radically new approach is required to find out what, if anything, MNs contribute to social behavior.

Theories relating to MN function have mainly been inspired by 'reflection' on the field properties of MNs found in a sample of laboratory monkeys with unreported (and possibly unknown) developmental histories. This method asks what neurons with these field properties might enable the animal to do. For example, early reports that MNs discharged when monkeys saw and produced object-directed actions inspired the theory that MNs mediate action understanding' via 'motor resonance'. Even now, opposition to the idea that MNs mediate action understanding tends to be answered by focusing on the conditions in which they fire (Gallese et al., 2011). The associative account suggests that the 'reflection' method needs to be changed and extended by embedding mirror neuron research in system-level theories of social behavior, by considering individuals' developmental history, and by carrying out experimental investigation of mirror neuron function.

If MNs were an adaptation, one could argue that new categories of psychological functioning – such as 'action understanding' and 'motor resonance' - are necessary to characterize what they do. In contrast, by showing that established psychological theory – associative learning theory - can cast

light on the origin of MNs, the associative account underlines the value of embedding research on MN function within system-level psychological and computational theories of how the brain produces behavior. This implies that hypotheses about MN function should specify a part in a process – a process that goes all the way from peripheral sensory input to overt motor output - that MNs are thought to fulfill. The name assigned to this part is not important in itself. What is important is that the hypothetical function of MNs is distinguished clearly from other components of the same overall process. For example, in this kind of system-level, theory-guided approach, ‘action understanding’ would be distinguished from components that are more purely perceptual (which might be called ‘action perception’ or ‘action recognition’), more purely motoric (e.g. ‘action execution’), or constitute a higher level of ‘understanding’ (e.g. mentalising). This approach would also make it clear whether the hypothetical function is thought to be optional or obligatory; whether it can be, or must be, done by MNs. The kind of system-level theoretical approach required in research on the functions of MNs is exemplified by studies of their role in speech perception (Lotto, Hickok, & Holt, 2009; Scott, McGettigan, & Eisner, 2009).

Regarding MN development, if MNs were an adaptation, it is likely that their properties would be relatively invariant across developmental environments. Therefore, it would be possible to make valid inferences about species-typical properties of MNs based on a relatively small and developmentally atypical sample of individuals. If MNs are instead a product of associative learning, this kind of inference is not valid. Whether or not an individual has MNs, which actions are encoded by their MNs, and at what level of abstraction, will all depend on the types of sensorimotor experience received by the individual in the course of their development. Therefore, the associative account implies that it is crucial for studies of laboratory monkeys to report, and ideally to control, the animals’ developmental history; the kinds of sensorimotor experience to which they have been exposed. A corollary of this is that we cannot assume that the mirror mechanisms found in the members of one human culture are representative of the whole human species. With its emphasis

on the role of social practices in driving the development of MNs, the associative account provides specific, theory-driven motivation for cross-cultural studies of mirroring.

In terms of function, a system-level theoretical approach would overcome a problem that has haunted discussions of the action understanding hypothesis since MNs were discovered: Is this hypothesis claiming that MN activity causes or constitutes action understanding? The former is an empirically testable hypothesis suggesting that there is a distinctive behavioral competence, called action understanding, to which the activity of MNs contributes. The latter implies that the firing of MNs during action observation is, in itself, a form of action understanding; it does not need to have further consequences in order to qualify as action understanding. This claim is not subject to empirical evaluation; it is true, or otherwise, by virtue of the meanings of words.

Empirical (rather than constitutive) claims about the function of MNs need to be tested by experiments looking for, at minimum, covariation between MN activity and behavioral competence, and, ideally, testing for effects on behavioral competence of interventions that change MN activity. At present, this research faces two major challenges. First, because the hypothetical functions of MNs typically are not defined in the context of a system-level theory, it is difficult to design appropriate control tasks. For example, if an experiment is testing the hypothesis that MNs play a causal role in action understanding, should it control for the possibility that they instead play some role in action perception? If so, what kind of behavioral competence is indicative of action perception rather than action understanding? To date, only a small number of studies (e.g. Pobric & Hamilton, 2006) have made a serious attempt to tackle this problem. The second challenge is that, with rare exceptions (Mukamel et al., 2010), MN activity cannot be localized precisely within the human brain. Consequently, many studies assume that activity in the ventral PMC and IPL – areas homologous to those in which MNs have been found in macaques – is MN activity, and that behavioral changes brought about through interference with the functioning of these areas are due

to interference with MNs. This is unsatisfactory because, in macaques, it is likely that fewer than 20% of the neurons in these classical mirror areas are actually MNs, and because there is evidence of MNs in nonclassical areas in both macaques and humans (see first section). Techniques such as fMRI repetition suppression, MVPA, and TMS adaptation (Cattaneo et al., 2011; Silvanto et al., 2007) hold some promise as means of overcoming the localization problem with human participants, by isolating behavioral effects to specific populations of neurons. Guided by system-level theory, future studies could use these techniques with a range of tasks to isolate the processes in which MNs are involved.

Alongside such future studies with human participants, animal studies could be conducted, not only to document the field properties of MNs, but to examine how those properties relate to behavioral competence. For example, are animals with MNs for actions X and Y better than other animals of the same species at behavioral discrimination of X and Y, or at imitating X and Y? Studies of this kind have been dismissed as impractical on the assumption that they would have to involve monkeys, and that between-group variation in MN activity would have to be induced via lesions or disruptive TMS. However, the associative account suggests that between-group variation in the number and type of MNs could be induced using sensorimotor training, either in monkeys or by establishing a rodent model. If the associative account is correct, rodents, birds, and other animals are likely to have the potential to develop MNs because they are capable of associative learning. Whether or not they receive in the course of typical development the sensorimotor experience necessary to realize this potential, relevant sensorimotor training could be provided in the laboratory.

<A>Conclusion

The associative account of the origin of MN properties paves the way for an alternative approach to MN research. It acknowledges that MNs were a fascinating discovery, and is open to the possibility

that they play one or more important roles in social interaction. It differs from the adaptation account in suggesting that sensorimotor learning plays a crucial, inductive role in the development of MNs, and, because of this, we will get reliable information about the function of MNs only by applying an approach based on system-level theory, developmental history, and experimentation. These methodological implications underline the fact that, relative to the adaptation account, the associative account shifts the balance of explanatory power from MNs themselves to the environments in which they develop. In some ways this is inconvenient because developmental environments are much harder to study in the laboratory, but there are significant potential payoffs. As a rich source of testable predictions about when, where and how MNs develop, associative learning theory can provide clear guidance for future research on the taxonomic distribution, typical properties and functional roles of MNs.

ACKNOWLEDGEMENTS:

CC is supported by the ESRC (ES/K00140X/1). We are also very grateful to Richard Cook and Geoff Bird with whom we developed the associative account of mirror neurons.

REFERENCES:

- Anisfeld, M. (1996). Only tongue protrusion modeling is matched by neonates. *Developmental Review*, 16(2), 149-161. 10.1006/drev.1996.0006
- Ariew, A. (2006). Innateness. In M. Matthen & C. Stevens (Eds.), *Handbook of the philosophy of science* (Vol. 3, pp. 567-584). Oxford: Elsevier.
- Baker, A. G., Vallée-Tourangeau, F., & Murphy, R. A. (2000). Asymptotic judgment of cause in a relative validity paradigm. *Memory & Cognition*, 28(3), 466-479.
- Blakesee, S. (2006). Cells that read minds. *The New York Times*, January 10.
- Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., & Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cerebral Cortex*, 20(6), 1372-1385. 10.1093/cercor/bhp200
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114(1), 80-99. 10.1037/0033-2909.114.1.80
- Bouton, M. E. (1994). Context, ambiguity, and classical-conditioning. *Current Directions in Psychological Science*, 3(2), 49-53. 10.1111/1467-8721.ep10769943
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 106(1-2), 3-22. 10.1016/S0001-6918(00)00024-X,

- Caggiano, V., Fogassi, L., Rizzolatti, G., Casile, A., Giese, M. A., & Thier, P. (2012). Mirror neurons encode the subjective value of an observed action. *Proceedings of the National Academy of Sciences of the United States of America*, 109(29), 11848-11853. 10.1073/pnas.1205553109
- Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., et al. (2011). View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Current Biology*, 21(2), 144-148. 10.1016/j.cub.2010.12.022,
- Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science*, 324(5925), 403-406. 10.1126/science.1166818
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, 15(8), 1243-1249. 10.1093/cercor/bhi007
- Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, 16(19), 1905-1910. 10.1016/j.cub.2006.07.065
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *European Journal of Neuroscience*, 28(6), 1208-1215. 10.1111/j.1460-9568.2008.06419.x
- Catmur, C., Mars, R. B., Rushworth, M. F., & Heyes, C. (2011). Making mirrors: Premotor cortex stimulation enhances mirror and counter-mirror motor facilitation. *Journal of Cognitive Neuroscience*, 23, 2352-2362. 10.1162/jocn.2010.21590

Catmur, C., Press, C., Cook, R., Bird, G. & Heyes, C. M, (in press) Mirror neurons: tests and testability. *Behavioral and Brain Sciences*.

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology*, 17(17), 1527-1531. 10.1016/j.cub.2007.08.006

Cattaneo, L., Barchiesi, G., Tabarelli, D., Arfeller, C., Sato, M., & Glenberg, A. M. (2011). One's motor performance predictably modulates the understanding of others' actions through adaptation of premotor visuo-motor neurons. *Social Cognitive and Affective Neuroscience*, 6(3), 301-310. 10.1093/scan/nsq099

Cavallo, A., Heyes, C., Becchio, C., Bird, G., & Catmur, C. (2013). Timecourse of mirror and counter-mirror effects measured with transcranial magnetic stimulation. *Social Cognitive and Affective Neuroscience*, in press. 10.1093/scan/nst085

Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Books.

Chong, T. T. J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology*, 18(20), 1576-1580. 10.1016/j.cub.2008.08.068

Coll, M.-P., Bird, G., Catmur, C., & Press, C. (submitted). Crossmodal repetition effects in the alpha oscillatory signal demonstrate tactile processing during action observation.

Cook, R. (2012). The ontogenetic origins of mirror neurons: evidence from 'tool-use' and 'audiovisual' mirror neurons. *Biology Letters*, 8(5), 856-9. 10.1098/rsbl.2012.0192

- Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C.M. (in press). Mirror neurons: from origin to function. *Behavioral and Brain Sciences*.
- Cook, R., Dickinson, A., & Heyes, C. (2012). Contextual modulation of mirror and countermirror sensorimotor associations. *Journal of Experimental Psychology: General*, 141(4), 774-87. 10.1037/a0027561
- Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance*, 36(4), 840-852. 10.1037/a0019256
- Cooper, R., Cook, R., Dickinson, A., & Heyes, C. (2013). Associative (not Hebbian) learning and the mirror neuron system. *Neuroscience Letters*, 540, 28-36. 10.1016/j.neulet.2012.10.002
- Cosmides, L., & Tooby, J. (1994). Beyond intuition and instinct blindness: toward an evolutionary rigorous cognitive science. *Cognition*, 50, 41-77. 10.1016/0010-0277(94)90020-5
- D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, 24(3), 955-958. 10.1111/j.1460-9568.2006.04960.x
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91(1), 176-180.
- Dushanova, J., & Donoghue, J. (2010). Neurons in primary motor cortex engaged during action observation. *European Journal of Neuroscience*, 31(2), 386-398. 10.1111/j.1460-9568.2009.07067.x

- Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, 68(2-3), 138-154. 10.1007/s00426-003-0151-8
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73(6), 2608-2611.
- Ferrari, P. F., Bonini, L., & Fogassi, L. (2009). From monkey mirror neurons to primate behaviours: possible 'direct' and 'indirect' pathways. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 364(1528), 2311-2323. 10.1098/rstb.2009.0062
- Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, 17(8), 1703-1714. 10.1046/j.1460-9568.2003.02601.x
- Ferrari, P. F., Rozzi, S., & Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience*, 17(2), 212-226. 10.1162/0898929053124910
- Ferrari, P. F., Vanderwert, R. E., Paukner, A., Bower, S., Suomi, S. J., & Fox, N. A. (2012). Distinct EEG amplitude suppression to facial gestures as evidence for a mirror mechanism in newborn monkeys. *Journal of Cognitive Neuroscience*, 24(5), 1165-1172. 10.1162/jocn_a_00198
- Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A., & Suomi, S. J. (2006). Neonatal imitation in rhesus macaques. *PLoS Biology*, 4(e302). 10.1371/journal.pbio.0040302

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722), 662-667.

10.1126/science.1106138

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119 (Pt 2), 593-609. 10.1093/brain/119.2.593

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance XIX*. Oxford: Oxford University Press.

Gallese, V., Gernsbacher, M., Hickok, G., Heyes, C., & Iacoboni, M. (2011). Mirror neuron forum. *Perspectives on Psychological Science*, 6(4), 369-407. 10.1177/1745691611413392

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501. 10.1016/S1364-6613(98)01262-5

Gallese, V., Rochat, M., Cossu, G., & Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Developmental Psychology*, 45(1), 103-113. 10.1037/a0014436

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences*, 15(11), 512-519. 10.1016/j.tics.2011.09.003

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, 19(6), 1239-1255. 10.1093/cercor/bhn181

Gillmeister, H., Catmur, C., Liepelt, R., Brass, M., & Heyes, C. (2008). Experience-based priming of body parts: A study of action imitation. *Brain Research*, 1217, 157-170.

10.1016/j.brainres.2007.12.076

Glenberg, A. (2011). Introduction to the Mirror Neuron Forum. *Perspectives on Psychological Science*, 8(4), 363-368. 10.1177/1745691611412386

Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Neural and behavioral plasticity* (pp. 24-54). New York: Academic Press.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14-23.

10.1016/j.tics.2005.11.006

Hari, R., & Salmelin, R. (1997). Human cortical oscillations: A view through the skull. *Trends in Neuroscience*, 20(1), 44-49. 10.1016/S0166-2236(96)10065-5

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, 34(4), 575-583. 10.1016/j.neubiorev.2009.11.007,

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, 137(3), 463-483. 10.1037/a0022288

Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society B- Biological Sciences*, 367(1603), 2695-703.

10.1098/rstb.2012.0217

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation.

Cognitive Brain Research, 22(2), 233-240. 10.1016/j.cogbrainres.2004.09.009

Heyes, C. M. & Ray, E. (2000). What is the significance of imitation in animals? *Advances in the Study of Behavior*, 29, 215-245.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7), 1229-1243. 10.1162/jocn.2009.21189

Iacoboni, M. (2008). The mirror neuron revolution: Explaining what makes humans social. *Scientific American*, July 1.

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, 60, 653-670. 10.1146/annurev.psych.60.110707.163604

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286(5449), 2526-2528.
10.1126/science.286.5449.2526

Jeannerod, M. (1994). The representing brain. Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187-245. 10.1017/S0140525X00034026

Keysers, C., & Gazzola, V. (2010). Social neuroscience: mirror neurons recorded in humans. *Current Biology*, 20(8), R353-354. 10.1016/j.cub.2010.03.013

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153(4), 628-636.

10.1007/s00221-003-1603-5

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, 29(32), 10153-10159. 10.1523/

JNEUROSCI.2668-09.2009

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, 13(6), 522-525. 10.1016/S0960-9822(03)00165-9

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582), 846-

848. 10.1126/science.1070311

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for

action suppression? *Neuron*, 64(6), 922-930. 10.1016/j.neuron.2009.12.010

Landmann, C., Landi, S. M., Grafton, S. T., & Della-Maggiore, V. (2011). fMRI supports the sensorimotor theory of motor resonance. *PLoS One*, 6(11), 1-8.

10.1371/journal.pone.0026859

Lepage, J. F., & Theoret, H. (2007). The mirror neuron system: grasping others' actions from birth?

Developmental Science, 10(5), 513-523. 10.1111/j.1467-7687.2007.00631.x

- Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24), 9925-9930. 10.1073/pnas.0902262106
- Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13(3), 110-114. 10.1016/j.tics.2008.11.008
- Mameli, M., & Bateson, P. (2006). Innateness and the sciences. *Biology and Philosophy*, 21, 155-188. 10.1007/s10539-005-5144-0
- Marshall, P. J., & Meltzoff, A. N. (2011). Neural mirroring systems: Exploring the EEG mu rhythm in human infancy. *Developmental Cognitive Neuroscience*, 1, 110-123. 10.1016/j.dcn.2010.09.001
- Marshall, P. J., Young, T., & Meltzoff, A. N. (2011). Neural correlates of action observation and execution in 14-month-old infants: An event-related EEG desynchronization study. *Developmental Science*, 14, 474-480. 10.1111/j.1467-7687.2010.00991.x
- Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neuroscience and Biobehavioral Reviews*, 36(1), 341-349. 10.1016/j.neubiorev.2011.07.004
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, 20(8), 750-756. 10.1016/j.cub.2010.02.045

- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, 78(4), 2226-2230.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology*, 83(5), 2580-2601.
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, 10(7), 817-818. 10.1038/nn1911
- Nyström, P. (2008). The infant mirror neuron system studied with high density EEG. *Social Neuroscience*, 3, 334-347. 10.1080/17470910701563665
- Oosterhof, N.N., Tipper, S.P., & Downing, P.E. (2012). Viewpoint (in)dependence of action representations: an MVPA study. *Journal of Cognitive Neuroscience*, 24(4), 975-989. 10.1162/jocn_a_00195
- Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to biological motion stimuli. *Journal of Cognitive Neuroscience*, 6, 99-116. 10.1162/jocn.1994.6.2.99
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, 76(1), 109-129.

- Paukner, A., Ferrari, P. F., & Suomi, S. J. (2011). Delayed imitation of lipsmacking gestures by infant rhesus macaques (*Macaca mulatta*). *PLoS ONE*, 6(12), e28848.
10.1371/journal.pone.0028848
- Paulus, M., Hunnius, S., van Elk, M., & Bekkering, H. (2012). How learning to shake a rattle affects 8-month-old infants' perception of the rattle's sound: Electrophysiological evidence for action-effect binding in infancy. *Developmental Cognitive Neuroscience*, 2, 90-96.
10.1016/j.dcn.2011.05.006
- Peck, C. A., & Bouton, M. E. (1990). Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learning and Motivation*, 21(1), 1-31. 10.1016/0023-9690(90)90002-6
- Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, 146, 87-113.
- Petroni, A., Baguear, F., & Della-Maggiore, V. (2010). Motor resonance may originate from sensorimotor experience. *Journal of Neurophysiology*, 104(4), 1867-1871. 10.1152/jn.00386.2010
- Pinker, S. (1997). *How the Mind Works*. London: The Penguin Press.
- Pobric, G., & Hamilton, A. F. (2006). Action understanding requires the left inferior frontal cortex. *Current Biology*, 16(5), 524-529. 10.1016/j.cub.2006.01.033

- Press, C., Catmur, C., Cook, R., Widmann, H., Heyes, C., & Bird, G. (2012). fMRI evidence of 'mirror' responses to geometric shapes. *PLoS One*, 7(12), e51934. 10.1371/journal.pone.0051934
- Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B-Biological Sciences*, 274(1625), 2509-2514. 10.1098/rspb.2007.0774
- Ramachandran, V. S. (2009). The neurons that shaped civilization. *TED*, from http://www.ted.com/talks/vs_ramachandran_the_neurons_that_shaped_civilization.html
- Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental Science*, 14(1), 92-105. 10.1111/j.1467-7687.2010.00961.x
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66(1), 1-5.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5), 188-194. 10.1016/S0166-2236(98)01260-0
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71(3), 491-507.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192. 10.1146/annurev.neuro.27.070203.144230

- Rizzolatti, G., & Fadiga, L. (1998). Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). *Novartis Foundation Symposium*, 218, 81-95; discussion 95-103.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2), 131-141. 10.1016/0926-6410(95)00038-0
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11(4), 264-274. 10.1038/nrn2805
- Rochat, M. J., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., et al. (2010). Responses of mirror neurons in area F5 to hand and tool grasping observation. *Experimental Brain Research*, 204(4), 605-616. 10.1007/s00221-010-2329-9
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, 23, 473-500. 10.1146/annurev.neuro.23.1.473
- Scott, S. K., McGettigan, C., & Eisner, F. (2009). A little more conversation, a little less action-candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, 10(4), 295-302. 10.1038/nrn2603
- Silvanto, J., Muggleton, N. G., Cowey, A., & Walsh, V. (2007). Neural activation state determines behavioral susceptibility to modified theta burst transcranial magnetic stimulation. *European Journal of Neuroscience*, 26(2), 523-528. 10.1111/j.1460-9568.2007.05682.x

- Southgate, V., Johnson, M. H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infants' online prediction of others' goals. *Psychological Science*, 21, 355-359.
10.1177/0956797610362058
- Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6), 1746-1759. 10.1037//0096-1523.26.6.1746
- Tkach, D., Reimer, J., & Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *Journal of Neuroscience*, 27(48), 13241-13250. 10.1523/JNEUROSCI.2895-07.2007
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 31(1), 155-165. 10.1016/S0896-6273(01)00337-3
- Verschoor, S. A., Weidema, M., Biro, S., & Hommel, B. (2010). Where do action goals come from? Evidence for spontaneous action-effect binding in infants. *Frontiers in Psychology*, 1(201), 1:6. 10.3389/fpsyg.2010.00201
- Vogt, S., Buccino, G., Wohlschlager, A. M., Canessa, N., Shah, N. J., Zilles, K., et al. (2007). Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage*, 37(4), 1371-1383. 10.1016/j.neuroimage.2007.07.005
- Wiggett, A. J., Hudson, M., Tipper, S. P., & Downing, P. E. (2011). Learning associations between action and perception: effects of incompatible training on body part and spatial priming. *Brain and Cognition*, 76(1), 87-96. 10.1016/j.bandc.2011.02.014

Williams, G. C. (1966). *Adaptation and Natural Selection*. New Jersey: Princeton University Press.

FIGURE CAPTIONS:

Figure 1: Mirror neuron areas in A) the macaque and B), C) the human brain. These are areas in which there is evidence at the single-cell or population level of neurons with sensorimotor matching properties. PMC, premotor cortex; IPL, inferior parietal lobule; IFG, inferior frontal gyrus.

Figure 2: Types of mirror neuron in the macaque. Typical sensory properties of four different types of sensorimotor neuron are shown; for simplicity, the same motor property (a precision grip) is shown for each mirror neuron type.

Figure 3: Mirror neurons from associative learning. A) Before learning, sensory neurons encoding visual descriptions of observed action are not systematically connected to motor neurons in parietal and premotor areas involved in the production of similar actions. B, a-d) Through social interaction and self-observation in the course of typical development, agents receive correlated sensorimotor experience; they see and do the same action at about the same time (contiguity), with one event predicting the other (contingency). This experience produces correlated activation of sensory and motor neurons coding similar actions, and, through associative learning, C) strengthens connections between these neurons. Due to these connections, neurons that were once involved only in the execution of action will also discharge during observation of a similar action; motor neurons become mirror neurons. Figure reproduced with permission from Heyes (2010).

Figure 4: Examples of contingencies which would produce A) strictly congruent, B) broadly congruent, and C) context-dependent ('grasp-to-place') mirror neurons. A) When a monkey watches its own actions while feeding, alone or in a group, the probability of seeing a particular grip (e.g., a precision grip) while performing exactly the same grip is high. B) When a monkey watches the actions of others during group feeding, the probability of seeing a range of grasping actions while performing a particular (e.g., precision) grip is also high (and, crucially, it is higher than the

probability of seeing an unrelated action e.g. a kick). C) Before testing for the presence of 'grasp-to-place' mirror neurons, monkeys are trained: i) when a pot is not present, food items should be eaten, but ii) when a pot is present, food items should be placed in the pot (in return for a higher-value food reward). Self-observation during this training ensures that, in the presence of a pot, the probability of seeing a grasp-to-place action while performing a grasp-to-place action is high. Subsequently, in the presence of a pot, the sight of a grasping action activates grasp-to-place (rather than grasp-to-eat) motor commands.

BIOGRAPHICAL NOTES:

Caroline Catmur is an ESRC Future Research Leader and Senior Lecturer in Cognitive Psychology at the University of Surrey, UK. Her research investigates the neural basis of social cognitive abilities including imitation, empathy and theory of mind.

Clare Press is a Lecturer in Psychological Sciences at Birkbeck, University of London. Her research examines the mechanisms underlying perception-action mapping for social processes such as imitation, as well as action selection and control functions.

Cecilia Heyes is a Senior Research Fellow in Theoretical Life Sciences and Professor of Psychology at All Souls College, University of Oxford, and a Fellow of the British Academy. Her work concerns the evolution of cognition; the ways in which natural selection, learning, developmental and cultural processes combine to produce the mature cognitive abilities found in adult humans.